OPIC
Office de la propriété intellectuelle du Canada

CIPO
Canadian Intellectual Property Office

Ottawa Hull K1A 0C9

(21) (A1) 2,196,496
(22) 1997/01/31
(43) 1998/08/01

(51) Int.Cl.[6] C12Q 1/25; G01N 33/68

(19)(CA) **APPLICATION FOR CANADIAN PATENT** (12)

(54) Protein Fragment Complementation Assay for the Detection of Protein-Protein Interactions

(72) Michnick, Stephen - Canada ;

(71) UNIVERSITÉ DE MONTRÉAL - Canada ;

(57) 1 Claim

ABSTRACT OF THE DISCLOSURE

We describe a strategy for designing protein-fragment complementation assays to detect protein-protein interactions *in vivo*. The design and implementation of this strategy is illustrated with the enzyme murine dihydrofolate reductase (DHFR). Fusion peptides consisting of *N*- and *C*-terminal fragments of murine DHFR fused to GCN4 leucine zipper sequences were coexpressed in *Escherichia coli* grown in minimal medium, where the endogenous DHFR activity was inhibited with trimethoprim. Coexpression of the complementary fusion products restored colony formation. Survival only occurred when both DHFR fragments were present and contained leucine-zipper forming sequences, demonstrating that reconstitution of enzyme activity requires assistance of leucine zipper formation. DHFR fragment-interface point mutants of increasing severity (Ile to Val, Ala and Gly) resulted in a sequential increase in *E. coli* doubling times illustrating the successful DHFR fragment reassembly rather than non-specific interactions between fragments. This *in vivo* assay could be used to study equilibrium and kinetic aspects of protein-protein interactions and for screening cDNA libraries for binding of a target protein with unknown proteins. Further, the selection and design criteria applied here to DHFR could be used to develop a strategy using any enzyme that would be appropriate for a clonal selection or direct detection of enzyme activity.

ABSTRACT OF THE DISCLOSURE

We describe a strategy for designing protein-fragment complementation assays to detect protein-protein interactions *in vivo*. The design and implementation of this strategy is illustrated with the enzyme murine dihydrofolate reductase (DHFR). Fusion peptides consisting of *N*- and C-terminal fragments of murine DHFR fused to GCN4 leucine zipper sequences were coexpressed in *Escherichia coli* grown in minimal medium, where the endogenous DHFR activity was inhibited with trimethoprim. Coexpression of the complementary fusion products restored colony formation. Survival only occurred when both DHFR fragments were present and contained leucine-zipper forming sequences, demonstrating that reconstitution of enzyme activity requires assistance of leucine zipper formation. DHFR fragment-interface point mutants of increasing severity (Ile to Val, Ala and Gly) resulted in a sequential increase in *E. coli* doubling times illustrating the successful DHFR fragment reassembly rather than non-specific interactions between fragments. This *in vivo* assay could be used to study equilibrium and kinetic aspects of protein-protein interactions and for screening cDNA libraries for binding of a target protein with unknown proteins. Further, the selection and design criteria applied here to DHFR could be used to develop a strategy using any enzyme that would be appropriate for a clonal selection or direct detection of enzyme activity.

TITLE OF THE INVENTION

A protein fragment complementation assay for the detection of protein-protein interactions

FIELD OF THE INVENTION

The present invention relates to the determination of the function of novel gene products. The invention further relates to a Protein fragment Complementation Assay (PCA). The PCA allows for the detection of a wide variety of types of protein-protein interactions in different contexts.

BACKGROUND OF THE INVENTION

Many processes in biology, including transcription, translation, and metabolic or signal transduction pathways, are mediated by non-covalently-associated multienzyme complexes[1-10]. Further, with rapid advances in genome sequencing projects there is a need to develop strategies to define "protein linkage maps", detailed inventories of protein interactions that make up functional assemblies of proteins[2,3]. Despite the importance of understanding protein assembly in biological processes, there are few convenient methods for studying protein-protein interactions *in vivo* [4,5]. A powerful and commonly used strategy, the yeast two-hybrid system, is used to identify novel protein-protein interactions and to examine the amino acid determinants of specific protein interactions[4,6-8]. The approach allows for screening of cDNA libraries, or mutants of individual genes. Cellular machineries for transcription, translation, and

metabolic or signal transduction pathways are examples of processes mediated by multiprotein complexes. The formation of multiprotein complexes produce the most efficient chemical machinery.

Much of modern biological research is concerned with identifying proteins involved in cellular processes, determining their functions and how, when, and where they interact with other proteins involved in specific pathways. Despite the importance of understanding protein assembly to biological processes, there are few available methods for studying protein-protein interactions *in vivo*. Approaches include the use of chemical crosslinking reagents and resonance energy transfer between dye-coupled proteins [102, 103]. A very powerful approach for studying protein-protein interactions *in vivo* is the recently developed yeast two-hybrid strategy [104-107]. This method is being used extensively to identify novel protein-protein interactions. The approach allows for rapid screening of a large number of clones, including cDNA libraries. Limitations of this technique include the fact that the interaction must occur in a specific context (the nucleus of *S. cerevisiae*), and generally cannot be used to distinguish induced *versus* constitutive interactions.

Recently, a novel strategy for detecting protein-protein interactions has been demonstrated by Johnsson and Varshavsky [108] called the ubiquitin-based split protein sensor (USPS)[9]. The strategy is based on cleavage of proteins with *N*-terminal fusions to ubiquitin by cytosolic proteases (ubiquitinases) that recognize its tertiary structure. The strategy depends on the reassembly of the tertiary structure of the protein ubiquitin from complementary *N*- and *C*-terminal fragments and crucially, on the augmention of this reassembly by oligomerization domains fused to these fragments. Reassembly is detected as specific

proteolysis of the assembled product by cytosolic proteases (ubiquitinases). The authors demonstrated that a fusion of a reporter protein-ubiquitin C-terminal fragment could also be cleaved by ubiquitinases, but only if co-expressed with an *N*-terminal fragment of ubiquitin that was complementary to the C-terminal fragment. The reconstitution of observable ubiquitinase activity only occurred if the *N*- and C-terminal fragments were bound through GCN4 leucine zippers [108,110]. The authors suggested that this "split-gene" strategy could be used as an *in vivo* assay of protein-protein interactions and analysis of protein assembly kinetics in cells. Unfortunately, this strategy requires additional agents (in this case ubiquitinases). There is therefore a need for a system which uses the reconstitution of a specific enzyme activity from fragments as the assay itself.

The methods described above require additional cellular machinery for detection that exist only in specific cellular compartments. There thus remains a need for the obtention of an oligomerization-assisted complementation of fragments which could provide a general strategy for detecting protein-protein interactions without a requirement for other proteins for the detection of the activity. Preferably, the assay would involve an oligomerization-assisted complementation of fragments of monomeric enzymes that require no other proteins for the detection of their activity. Furthermore, if the structure of an enzyme were known it would be possible to design fragments of the enzyme to ensure that the reassembled fragments would be active and to introduce mutations to alter the stringency of detection of reassembly. The flexibility allowed in the design of such an approach would make it applicable to situations where other detection systems may not be suitable.

Recent advances in human genomics research has led to rapid progress in the identification of novel genes. In applications to biological and pharmaceutical research, there is now the pressing need to determine the functions of novel gene products; for example, for genes shown to be involved in disease phenotypes. It is in addressing questions of function where genomics-based pharmaceutical research becomes bogged down and there is now the need for advances in the development of simple and automatable functional assays. A first step in defining the function of a novel gene is to determine its interactions with other gene products in an appropriate context; that is, since proteins make specific interactions with other proteins as part of functional assemblies, an appropriate way to examine the function of a novel gene is to determine its physical relationships with the products of other genes.

Screening techniques for protein interactions, such as the yeast "two-hybrid" system, have transformed molecular biology, but can only be used to study specific types of interactions and to screen for constitutively interacting proteins. To rationally screen for protein interactions within the context of a specific problem requires more flexible approaches.

There thus remains a need to develop a protein fragment complementation assay having the following capabilities:

1) Allow for the detection of protein-protein interactions *in vivo* or *in vitro*.

2) Allow for the detection of protein-protein interactions in appropriate contexts, such as within a specific organism, cell type, cellular compartment, or organelle.

3) Allow for the detection of induced *versus* constitutive protein-protein interactions (such as by a cell growth or inhibitory factor).

4) To be able to distinguish specific-*versus* non-specific protein-protein interactions by controlling the sensitivity of the assay.

5) Allow for the detection of the kinetics of protein assembly in cells.

6) Allow for screening of cDNA libraries for protein-protein interactions.

SUMMARY OF THE INVENTION

The present invention seeks to provide these and other needs. The present invention surprisingly provides an oligomerization-assisted complementation of fragments which provides a general strategy for detecting protein-protein. In a preferred embodiment, the instant invention seeks to provide an oligomerization-assisted complementation of fragments of monomeric enzymes that require no other proteins for the detection of their activity. In one such embodiment, a protein-fragment complementation assay (PCA) based on reconstitution of dihydrofolate reductase activity by complementation of defined fragments of the enzyme in *E. coli* is hereby provided. This assay requires no additional endogenous factors for detecting specific protein-protein interactions (i.e. leucine zipper interactions) and can be conveniently extended to screening cDNA libraries for protein interactions. In addition, the assay can also be adapted for detection of protein-protein interactions in any cellular context or compartment and be used to distinguish between induced *versus* constitutive protein-protein interactions in both prokaryotic and eukaryotic systems.

One particular strategy for designing a protein complementation assay (PCA) is based on using the following characteristics: 1) An enzyme that is relatively small and monomeric, 2) for which there is a large literature of structural and functional information,

3) for which simple assays for the activity of the enzyme both *in vivo* and *in vitro* exist, and 4) for which overexpression in eukaryotic and prokaryotic cells has been demonstrated. If these criteria are met, the structure of the enzyme is used to decide the best position in the polypeptide chain to split the gene in two, based on the following criteria: 1) The fragments should result in subdomains of continuous polypeptide; that is, the resulting fragments will not disrupt the subdomain structure of the protein, 2) the catalytic and cofactor binding sites should all be contained in one fragment, and 3) resulting new *N*- and *C*-termini should be on the same face of the protein to avoid the need for long peptide linkers and allow for studies of orientation-dependence of protein binding.

It should be understood that the above mentioned criteria do not all need to be satisfied for a proper working of the present invention. It is an advantage that the enzyme be small, preferably between 10-40kDa. Although monomeric enzymes are preferred, multimeric enzymes can also be envisaged as within the scope of the present invention. The dimeric protein tyrosinase can be used in the instant assay. The information on the enzyme provides an additional advantage in designing the PCA. Criteria 3) above is a crucial criteria. Although the overexpression in prokaryotic cells is preferred it is not an absolute necessity. It will be understood to the skilled artisan that the enzyme catalytic site (of the chosen enzyme) does not absolutely need to be on same molecule.

The present application explains the rationale and criteria for using a particular enzyme in a PCA. Figure A shows a general description of a PCA. Using molecular biology techniques, the chosen fragments of the enzyme are subcloned, and to the 5' ends of each,

proteins that either are known or thought to interact are fused. Co-transfection or transformation these DNA constructs into cells is then carried out and reconstitution with some assay is observed. It is crucial to understand that these assays will only work if the fused, interacting proteins catalyze the reassembly of the enzyme. That is, observation of reconstituted enzyme activity must be a measure of the interaction of the fused proteins. The gene for an enzyme is rationally dissected into two or more pieces . Fusion proteins are constructed with two proteins that are thought to bind to each other, fused to either of the two probe fragments. Reassembly of the probe protein from its fragments is catalyzed by the binding of the test proteins to each other, and is detected as reconstitution of enzyme activity.

A preferred embodiment of the present invention focuses on a PCA based on the enzyme dihydrofolate reductase. Expansion of the strategy to include assays in eukaryotic, cells, library screening, and a specific application to a problem concerning the role of the enzyme p70 S6 ribosomal protein kinase in integrating growth factor-activated signal transduction pathways is also provided. Four additional assays, including those based on the proteins Ubiquitin, Glutathione-S-Transferase, Green Fluorescent Protein, and Fire Fly Luciferase are also disclosed. Finally the present invention teaches how the PCA strategy can be both generalized and automated for functional testing of novel genes and the screening of natural products for pharmacological activity.

A simple strategy to design and implement assays for detecting protein-protein interactions *in vivo* is disclosed herein. We have designed complementary fragments of the native mDHFR that, when coexpressed in *E. coli* grown in minimal medium, allow for survival of

clones expressing the two fragments, where the basal activity of the endogenous bacterial DHFR is inhibited by the competitive inhibitor trimethoprim (Fig. 3). Reconstitution of activity only occurred when both *N*- and *C*-terminal fragments of DHFR were coexpressed as *C*-terminal fusions to GCN4 leucine zipper sequences, indicating that reassembly of the fragments requires formation of a leucine zipper between the *N*- and *C*-terminal fusion peptides. The sequential increase in cell doubling times resulting from the destabilizing mutations directed at the assembly interface (Ile114 to Val, Ala or Gly) demonstrates that the observed cell survival under selective conditions is a result of the specific, leucine-zipper-assisted association of mDHFR fragment[1,2] with fragment[3], as opposed to nonspecific interactions of Z-F[3] with Z-F[1,2].

As demonstrated previously with the ubiquitin-based split protein sensor (USPS)[9], a protein-fragment complementation strategy can be used to study equilibrium and kinetic aspects of protein-protein interactions *in vivo*. The DHFR PCA however, is a simpler assay based on the same strategy. It is a complete system; no additional endogenous factors are necessary and the results of complementation are observed directly, with no further manipulation. The *E. coli* cell survival assay described herein should therefore be particularly useful for screening cDNA libraries for protein-protein interactions. mDHFR expression in cells can be monitored by binding of fluorescent high-affinity substrate analogues for DHFR[26].

There are several further aspects of the DHFR PCA that distinguish it from all other strategies for studying protein-protein interactions *in vivo* (except USPS). We have designed complementary fragments of DHFR that allow for controlling the stringency of the assay,

and could be used to obtain estimates of the kinetics and equilibrium constants for association of two proteins. The point mutations of the wild-type mDHFR Ile 114 to Val, Ala, or Gly alter the stringency of reconstitution of DHFR activity. For determining estimates of equilibrium and kinetic parameters for a specific protein-protein interaction, one could perform a series of DHFR PCA experiments with two proteins that interact with a known affinity, using the wild type or destabilizing mutant DHFR fragments. Comparison of cell growth rates in this model system with rates for a DHFR PCA using unknowns would give an estimate of the strength of the unknown interaction.

It should be understood that the present invention should not be limited to the DHFR PCA, as it is only one non-limiting embodiment of the protein complementation assay of the present invention. Moreover, the DHFR PCA should not be limited in the context in which it could be used. Constructs could be designed for targeting the mDHFR fusions to specific compartments in the cell by addition of signaling peptide sequences[27,28]. Induced *versus* constitutive protein-protein interactions could be distinguished by a eukaryotic version of the DHFR PCA, in the case of an interaction that is triggered by a biochemical event. Finally, the system could be adapted for use in screening for novel, induced protein-protein associations between a target protein and an expression library.

In conclusion, the present invention is pionneering as it is the first protein complementation assay displaying such a level of simplicity and versatility. The examplified embodiment is a protein-fragment complementation assay (PCA) based on mDHFR, where a leucine zipper directs the reconstitution of DHFR activity. Activity was

detected by an *E. coli* survival assay which is both practical and inexpensive. This system illustrates the use of mDHFR fragment complementation in the detection of leucine zipper dimerization and could be applied to the detection of unknown, specific protein-protein interactions *in vivo*.

It should be undertstood that the instant invention is not limited to DHFR, as numerous other selectable markers can be selected and used in accordance with the teachings of the present invention. Examples of such markers can be found in Kaufman, (1987 Genetic Eng. 9:155-198) and references found therein. Non-limiting examples of selectable markers which satisfy the criteria as set forth herein include: Thymidine kinase, ribonucleotidereductase, glutamin synthetase, IMP-5' dehydrogenase, adenylate deaminase, methionine-1, and threonyl-tRNA synthetase.

It should also be clear to the skilled artisan to which the present invention pertains that the invention is not limited to the use of leucine zippers as the two protein-protein interaction domains. Indeed, numerous other types of protein-protein interactions can be used and identified in accordance with the teaching of the present invention. The known types of motifs involved in protein-protein interactions are well known in the art. Non-limiting examples of such domains include coiled-coil motifs, helix-loop-helix motifs.

The present application refers to a number of documents, the contents of which is herein incorporated by reference.

Other features and advantages of the present invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. a Figure A shows a general description of a PCA. Using molecular biology techniques, the chosen fragments of the enzyme are subcloned, and to the 5' ends of each, proteins that either are known or thought to interact are fused. Co-transfection or transformation these DNA constructs into cells is then carried out and reconstitution with some assay is observed.

FIG. 1 (A) A model of the Z-F[1,2] - Z-F[3] complex. Z-F[1,2] is represented in orange, blue, and cyan for the first helix of the GCN4 leucine zipper, F[1] and F[2] respectively. Z-F[3] is rendered in green and yellow for the second helix of the zipper and F[3] respectively. The model was derived from crystallographic coordinates for a dimeric leucine zipper of GCN4 bound to DNA [33] and the DHFR structure is from crystal structure coordinates for hDHFR complexed with folate[11].
(B) Structure of hDHFR bound to NADPH (violet) and methotrexate-g-tetrazole (gold)[34]. F [1,2] and [3] are colored as in (A). Ile 114 and surrounding residues in F[1,2] are shown in grey.

FIG. 2: Scheme of the fusion constructs used in this study. The hexahistidine peptide (6His), the homodimerizing GCN4 leucine zipper (Zipper) and mDHFR fragments (1, 2 and 3) are illustrated. The labels for the constructs are used to identify both the DNA constructs and the proteins expressed from these constructs.

FIG. 3: (A) *E. coli* survival assay on minimal medium plates. Control: Left side of the plate: *E. coli* harboring pQE-30 (no insert); right side: *E. coli* harboring pQE-16, coding for native mDHFR. Panel I: Left side of each plate: transformation with construct Z-F[1,2]; right side of each plate: transformation with construct Z-F[3]. Panel II: Cotransformation with constructs Z-F[1,2] and Z-F[3]. Panel III: Cotransformation with constructs Control-F[1,2] and Z-F[3]. All plates contain 0.5 mg/ml trimethoprim. In panels I to III, plates on the right side contain 1mM IPTG.

(B) *E. coli* survival assay using destabilizing DHFR mutants. Panel I: Cotransformation of *E. coli* with constructs Z-F[1,2] and Z-F[3:Ile114Val]. Panel II: Cotransformation with Z-F[1,2] and Z-F[3:Ile114Ala]. Inset is a 5-fold enlargement of the right-side plate. Panel III: Cotransformation with Z-F[1,2] and Z-F[3:Ile114Gly]. All plates contain 0.5 mg/ml trimethoprim. Plates on the right side contain 1mM IPTG.

FIG. 4: Coexpression of mDHFR fragments. (A) Agarose gel analysis of restriction pattern resulting from *Hinc*II digestion of plasmid DNA. Lane 1 contains DNA isolated from *E. coli* cotransformed with constructs Z-F[1,2] and Z-F[3]. Lanes 2 and 3 contain DNA isolated from *E. coli* transformed with, respectively, construct Z-F[3] and construct Z-F[1,2]. Fragment migration (in bp) is indicated to the right.

(B) SDS-PAGE analysis of mDHFR fragment expression. Lanes 1 to 5 show crude lysate of untransformed *E. coli* (lane 1), or *E. coli* expressing Z-F[1,2] (20.8 kDa; lane 2), Z-F[3] (18.4 kDa; lane 3), Control-F[1,2] (14.2 kDa; lane 4), and Z-F[1,2] + Z-F[3] (lane 5). Lane 6 shows 40 ml out of 2ml copurified Z-F[1,2] and Z-F[3].

Arrowheads point to the proteins of interest. Migration of molecular weight markers (in kDa) is indicated to the right.

Other objects, advantages and features of the present invention will become more apparent upon reading of the following non-restrictive description of preferred embodiments with reference to the accompanying drawings which are examplary and should not be interpreted as limiting the scope of the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENT

Selection of mDHFR for a PCA. Murine DHFR (mDHFR) meets all of the criteria for a PCA listed above. Prokaryotic and eukaryotic DHFR is central to cellular one-carbon metabolism and is absolutely required for cell survival in both prokaryotes and eukaryotes. Specifically it catalyses the reduction of dihydrofolate to tetrahydrofolate for use in transfer of one-carbon units required for biosynthesis of serine, methionine, purines and thymidylate. The DHFRs are small (17 kD to 21 kD), monomeric proteins. The crystal structures of DHFR from various bacterial and eukaryotic sources are known and substrate binding sites and active site residues have been determined [111-114], allowing for rational design of protein fragments. The folding, catalysis, and kinetics of a number of DHFRs have been studied extensively[115-119]. The enzyme activity can be monitored *in vitro* by a simple spectrophotometric assay [120], or *in vivo* by cell survival in cells grown in the absence of DHFR end products. DHFR is specifically inhibited by the anti-folate drug trimethoprim. As mammalian DHFR has a 12000-fold lower affinity for trimethoprim than does bacterial DHFR[121], growth of bacteria expressing

mDHFR in the presence of trimethoprim levels lethal to bacteria is an efficient means of selecting for reassembly of mDHFR fragments into active enzyme. High level expression of mDHFR has been demonstrated in transformed prokaryote or transfected eukaryotic cells[122-126].

Design Considerations. mDHFR shares high sequence identity with the human DHFR (hDHFR) sequence (91% identity) and is highly homologous to the *E. coli* enzyme (29% identity, 68% homology) and these sequences share visually superimposable tertiary structure [111]. Comparison of the crystal structures of mDHFR and hDHFR suggests that their active sites are essentially identical[127,128]. DHFR has been described as being formed of three structural fragments forming two domains[129, 130] the adenine binding domain (residues 47 to 105 = fragment[2]) and a discontinuous domain (residues 1 to 46 = fragment[1] and 106 to 186 [3]; numbering according to the murine sequence). The folate binding pocket and the NADPH binding groove are formed mainly by residues belonging to fragments[1] and [2]. Fragment [3] is not directly implicated in catalysis.

Residues 101 to 108 of hDHFR, at the junction between fragment[2] and fragment[3], form a disordered loop which lies on the same face of the protein as both termini. We chose to cleave mDHFR between fragments [1,2] and [3], at residue 107, so as to cause minimal disruption of the active site and NADPH cofactor binding sites. The native *N*- terminus of mDHFR and the novel *N*-terminus created by cleavage occur on the same surface of the enzyme[112, 126] allowing for ease of *N*-terminal covalent attachment of each fragment to associating fragments such as the leucine zippers used in this study. Using this

system, we have obtained leucine-zipper assisted assembly of the mDHFR fragments into active enzyme.

EXPERIMENTAL PROTOCOL

DNA Constructs. Mutagenic and sequencing oligonucleotides were purchased from Gibco BRL. Restriction endonucleases and DNA modifying enzymes were from Pharmacia and New England Biolabs. The mDHFR fragments carrying their own in-frame stop codon were subcloned into pQE-32 (Qiagen), downstream from and in-frame with the hexahistidine peptide and a GCN4 leucine zipper (Fig. 1A; Fig. 2). All final constructs were based on the Qiagen pQE series of vectors, which contain an inducible promoter-operator element (*tac*), a consensus ribosomal binding site, initiator codon and nucleotides coding for a hexahistidine peptide. Full-length mDHFR is expressed from pQE-16 (Qiagen).

Expression vector harboring the GCN4 leucine zipper: Residues 235 to 281 of the GCN4 leucine zipper (a *Sall/Bam*HI 254 bp fragment) were obtained from a yeast expression plasmid pRS316[9]. The recessed terminus at the *Bam*HI site was filled-in with Klenow polymerase and the fragment was ligated to pQE-32 linearized with *Sall/Hind*III(filled-in). The product, construct Z, carries an open reading frame coding for the sequence Met-Arg-Gly-Ser followed by a hexahistidine tag and 13 residues preceding the GCN4 leucine zipper residues.

Creation of DHFR fragments: The eukaryotic transient expression vector, pMT3 (derived from pMT2)[16], was used as a template for PCR-generation of mDHFR containing the features allowing subcloning and separate expression of fragment[1,2] and fragment[3]. The megaprimer method of

PCR mutagenesis[29] was used to generate a full-length 590 bp product. Oligonucleotides complementary to the nucleotide sequence coding for the *N*- and *C*-termini of mDHFR and containing a novel *Bsp*EI site outside the coding sequence were used as well as an oligonucleotide used to create a novel stop codon after fragment[1,2], followed by a novel *Spe*I site for use in subcloning fragment[3].

Construction of a new multiple cloning region and subcloning of DHFR fragments [1,2] and [3]: Complementary oligonucleotides containing the novel restriction sites: *Sna*BI, *Nhe*I, *Spe*I and *Bsp*EI, were hybridized together resulting in 5' and 3' overhangs complementary to *Eco*RI, and inserted into pMT3 at a unique *Eco*RI site. The 590 bp PCR product (described above) was digested with *Bsp*EI and inserted into pMT3 linearized at *Bsp*EI, yielding construct **[1,2,3]**. The 610 bp *Bsp*EI/*Eco*NI fragment (coding for DHFR fragment[1,2], followed by a novel stop and fragment[3] up to *Eco*NI) was filled in at *Eco*NI and subcloned into pMT3 opened with *Bsp*EI/*Hpa*I, yielding construct **F[1,2]**. The 250 bp *Spe*I/*Bsp*EI fragment of construct [1,2,3] coding for DHFR fragment[3] (with no in-frame stop codon) was subcloned into pMT3 opened with the same enzymes. The stop codon of the wild-type DHFR sequence, downstream from fragment[3] in pMT3, was inserted as follows. Cleavage with *Eco*NI, present in both the inserted fragment[3] and the wild-type fragment[3], removal of the 683 bp intervening sequence and religation of the vector yielded a construct of fragment[3] with the wild-type stop codon, construct **F[3]**.

Creation of the expression constructs: The 1051 bp and the 958 bp *Sna*BI/*Xba*I fragments of constructs F[1,2] and F[3], respectively, were subcloned into construct Z opened with *Bgl*I(filled-in)/*Nhe*I, yielding

constructs **Z-F[1,2]** and **Z-F[3]** (Fig. 1A; Fig. 2). For the Control expression construct, the 180 bp *XmaI/BspE*I fragment coding for the zipper was removed from construct Z-F[1,2], yielding construct **Control-F[1,2]** (Fig. 2).

Creation of Stability Mutants. Site-directed mutagenesis was performed[30] to produce mutants at Ile114 (numbering of the wild-type mDHFR). The mutagenesis reaction was carried out on the *KpnI/Bam*HI fragment of construct Z-F[3] subcloned into pBluescript SK$^+$ (Stratagene), using oligonucleotides that encode a silent mutation producing a novel *Bam*HI site. The 206 bp *NheI/Eco*NI fragment of putative mutants identified by restriction was subcloned back into Z-F[3]. The mutations were confirmed by DNA sequencing.

***E. coli* Survival Assay.** *E. coli* strain BL21 carrying plasmid pRep4 (from Qiagen, for constitutive expression of the *lac* repressor) were made competent, transformed with the appropriate DNA constructs and washed twice with minimal medium before plating on minimal medium plates containing 50 mg/ml kanamycin, 100 mg/ml ampicillin and 0.5 mg/ml trimethoprim. One half of each transformation mixture was plated in the absence, and the second half in the presence, of 1 mM IPTG. All plates were placed at 37°C for 66 hrs.

***E. coli* Growth Curves**: Colonies obtained from cotransformation were propagated and used to inoculate 10 ml of minimal medium supplemented with ampicillin, kanamycin as well as IPTG (1mM) and trimethoprim (1 $\mu g/\mu l$) where indicated. Cotransformants of Z-F[1,2] + Z-F[3:Ile114Gly] were obtained under non-selective conditions by plating the transformation mixture on L-agar (+ kanamycin and ampicillin) and screening for the presence of the two constructs by restriction analysis.

All growth curves were performed in triplicate. Aliquots were withdrawn periodically for measurement of optical density. Doubling time was calculated for early logarithmic growth (OD_{600} between 0.02 and 0.2).

Protein Overexpression and Purification. Bacteria were propagated in Terrific Broth[31] in the presence of the appropriate antibiotics to an OD_{600} of approximately 1.0. Expression was induced by addition of 1 mM IPTG and further incubation for 3 hrs. For analysis of crude extract, pellets from 150 ml of induced cells were lysed by boiling in loading dye. The lysates were clarified by microcentrifugation and analyzed by SDS-PAGE[32]. For protein purification, a cell pellet from 50 ml of induced *E. coli* cotransformed with constructs Z-F[1,2] and Z-F[3] was lysed by sonication, and a denaturing purification of the insoluble pellet undertaken using Ni-NTA (Qiagen) as described by the manufacturer. The proteins were eluted with a stepwise imidazole gradient. The fractions were analyzed by SDS-PAGE.

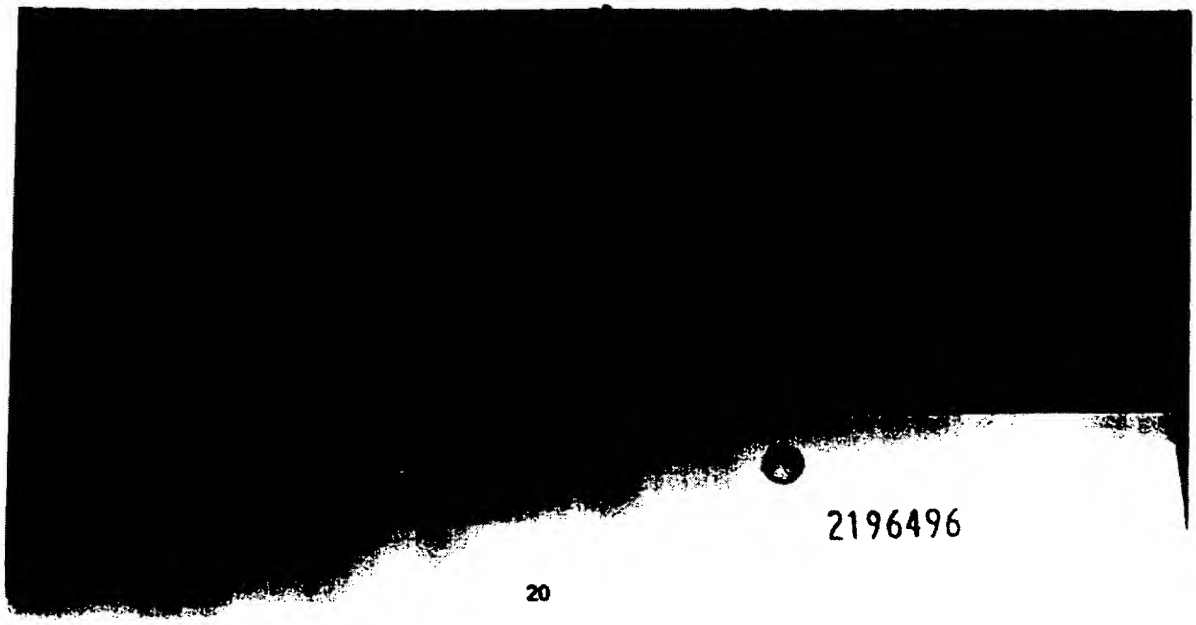
RESULTS

Selection of mDHFR for a PCA. In designing a protein-fragment complementation assay (PCA), we sought to identify an enzyme for which the following is true: 1) An enzyme that is relatively small and monomeric, 2) for which structural and functional information exists, 3) for which simple assays exist for both *in vivo* and *in vitro* measurement, and 4) for which overexpression in eukaryotic and prokaryotic cells has been demonstrated. Murine DHFR (mDHFR) meets all of these criteria. DHFR is central to prokaryotic and eukaryotic one-carbon metabolism and is absolutely required for cell survival. mDHFR is a small (21 kD), monomeric protein of known crystal structure[10-13]. The enzyme activity

can be monitored *in vivo* by cell survival in cells grown in the absence of DHFR end-products. In our assay we took advantage of the fact that *E. coli* DHFR is selectively inhibited by the anti-folate drug trimethoprim. As mammalian DHFR has a 12,000-fold lower affinity for trimethoprim than does bacterial DHFR[14], growth of bacteria expressing mDHFR in the presence of trimethoprim levels lethal to bacteria is an efficient means of selecting for reassembly of mDHFR fragments into an active enzyme. Finally, mDHFR expression has been demonstrated to occur at high levels in transformed prokaryotic or transfected eukaryotic cells[15, 16].

Design Considerations. mDHFR shares high sequence identity with the human DHFR (hDHFR) sequence (91% identity) and is highly homologous to the *E. coli* enzyme (29% identity, 68% homology). Comparison of the crystal structures of mDHFR and hDHFR suggests that their active sites and substrate binding pockets are identical, and homologous to those of *E. coli* DHFR[11, 17]. As the coordinates of the murine crystal structure were not available, we based our design considerations on the hDHFR structure[11], with further comparisons to the avian and *E. coli* structures[10, 13]. DHFR has been described as comprising three structural fragments forming two domains: the adenine binding domain (F[2]) and a discontinuous domain (F[1] and F[3])[13,18]. The folate binding pocket and the NADPH binding groove are formed mainly by residues belonging to F[1] and F[2]. Residues in F[3] contribute little to substrate binding and catalysis [11], but do contribute to DHFR stability and to kinetic parameters [19, 20].

Residues 101 to 108 of hDHFR, at the junction between F[2] and F[3], form a disordered loop which lies on the same face of the protein as both termini. Cleavage of mDHFR at this loop and fusion of



the native termini has produced a circularly permuted protein with physical and kinetic properties very similar to the native enzyme, suggesting that radical modifications in this loop are not disruptive to activity[21]. We chose to cleave mDHFR between F[1,2] and F[3], at residue 107, so as to cause minimal disruption of the active site and substrate binding sites. The native *N*-terminus of mDHFR and the novel *N*-terminus created by cleavage occur on the same surface of the enzyme[11,17] facilitating *N*-terminal covalent attachment of each fragment to associating fragments such as the leucine zippers used in this study (Fig. 1A).

***E. coli* Survival Assays.** Figure 2 illustrates the general features of the expressed constructs and the nomenclature used in this study. Figure 3 (panel A) illustrates the results of cotransformation of bacteria with constructs coding for Z-F[1,2] and Z-F[3], in the presence of trimethoprim, clearly showing that colony growth under selective pressure is possible only in cells expressing both fragments of mDHFR. There is no growth in the presence of either Z-F[1,2] or Z-F[3] alone. Induction of protein expression with IPTG is essential for colony growth (Fig. 3A). The presence of the leucine zipper on both fragments of mDHFR is essential as illustrated by cotransformation of bacteria with both vectors coding for mDHFR fragments, only one of which carries a leucine zipper (Fig. 3A). It should be noted that growth of control *E. coli* transformed with the full-length mDHFR is possible in the absence of IPTG due to low levels of expression in uninduced cells.

Confirmation of the presence of both plasmids in bacteria able to grow with trimethoprim was obtained from restriction analysis of the plasmid DNA purified from isolated colonies. Figure 4 (A)

reveals the presence of the 1200 bp *Hinc*II restriction fragment from construct Z-F[1,2] as well as the 487 and 599 bp *Hinc*II restriction fragments from construct Z-F[3]. Also present is the 935 bp *Hinc*II fragment of pRep4. Overexpression of the fusion proteins is illustrated in Figure 4 (B). In all cases, overexpression of a protein of the expected molecular weight is apparent on SDS-PAGE of the crude lysate. Purification of the coexpressed proteins under denaturing conditions yielded two bands of apparent homogeneity upon analysis by Coomassie-stained SDS-PAGE (Fig. 4B).

Stability Mutants: We generated mutants of F[3] to test whether reconstitution of mDHFR activity by fragment assembly was specific. Protein stability can be reduced by changing the side-chain volume in the hydrophobic core of a protein[9, 22-25]. Residue Ile114 of mDHFR (highlighted in Fig. 1B) occurs in a core b-strand at the interface between F[1,2] and F[3], isolated from the active site. Ile 114 is in van der Waals contact with Ile51 and Leu93 in F[1,2][11]. We mutated Ile 114 to Val, Ala, or Gly. Figure 3 (panel B) illustrates the results of cotransformation of *E. coli* with construct Z-F[1,2] and the mutated Z-F[3] constructs. The colonies obtained from cotransformation with Z-F[3:Ile114Ala] grew more slowly than those cotransformed with Z-F[3] or Z-F[3:Ile114Val] (see inset to Fig. 3B). No colony growth was detected in cells cotransformed with Z-F[3:Ile114Gly]. The number of transformants obtained was not significantly different in the case where colonies were observed, implying that cells cotransformed with Z-F[1,2] and either Z-F[3], Z-F[3:Ile114Val] or Z-F[3:Ile114Ala] have an equal survival rate. Overexpression of the mutants Z-F[3:Ile114X] was in the same range as Z-F[3], as determined by Coomassie-stained SDS-PAGE (data not shown).

We compared the relative efficiency of reassembly of mDHFR fragments by measuring the doubling time of the cotransformants in liquid medium. Doubling time in minimal medium was constant for all transformants (Table 1). Selective pressure by trimethoprim in the absence of IPTG prevented growth of *E. coli* except when transformed with pQE-16 coding for full-length DHFR due to low levels of expression in uninduced cells. Induction of mDHFR fragment expression with IPTG allowed survival of cotransformed cells (except in the case of Z-F[1,2] + Z-F[3:Ile114Gly], although the doubling times were significantly increased relative to growth in the absence of trimethoprim. The doubling time measured for cells expressing Z-F[1,2] + Z-F[3], Z-F[1,2] + Z-F[3:Ile114Val] and Z-F[1,2] + Z-F[3:Ile114Ala] were 1.6-fold, 1.9-fold and 4.1-fold, higher respectively, than the doubling time of *E. coli* expressing pQE-16 in the absence of trimethoprim and IPTG. The presence of IPTG unexpectedly prevented growth of *E. coli* transformed with full-length mDHFR. Growth was partially restored by addition of the folate metabolism end-products thymine, adenine, pantothenate, glycine and methionine (data not shown). This suggests that induced overexpression of mDHFR was lethal to *E. coli* when grown in minimal medium as a result of depletion of the folate pool by binding to the enzyme.

DISCUSSION

We are presently making point mutations in the GCN4 leucine zipper of Z-F[1,2] and Z-F[3], for which direct equilibrium and kinetic parameters are known and correlating these known values with parameters derived from the PCA (Pelletier and Michnick, *in preparation*). Comparison of cell growth rates in this model system with rates for a

DHFR PCA using unknowns would give an estimate of the strength of the unknown interaction. This should enable the determination of estimates of equilibrium and kinetic parameters for a specific protein-protein interaction.

In conclusion, we have developed and demonstrated a protein-fragment complementation assay (PCA) based on mDHFR, where a leucine zipper directs the reconstitution of DHFR activity. Activity was detected by an *E. coli* survival assay which is both practical and inexpensive. This system illustrates the use of mDHFR fragment complementation in the detection of leucine zipper dimerization and could be applied to the detection of unknown, specific protein-protein interactions *in vivo*.

EXAMPLE 1

THE mDHFR PCA assay

Application of the DHFR PCA to Mapping Growth Factor-Mediated Signal Transduction Pathways

One of the earliest detectable events in growth factor-activated cell proliferation is the serine phosphorylation of the S6 protein of the 40S ribosomal subunit. The discovery of serine/threonine kinases that specifically phosphorylate S6 have considerably aided in identifying novel mitogen mediated signal transduction pathways. The serine/threonine kinase $p70^{S6k}$ has been identified as a specific S6 phosphorylase[131-136]. $p70^{S6k}$ is activated by serine and threonine phosphorylation at specific sites in response to several mitogenic signals including serum in serum starved cells, growth factors including insulin and PDGF, and by mitogens such as phorbol esters. Considerable effort

has been made over the last five years to determine how p70/p85^{S6k} are activated in response to mitogens. Two receptor-mediated pathways have been implicated in p70^{S6k} activation, one associated with the phosphatidylinositol-3-kinase (PI(3)k) and the other with the PI(3)k homologue mTOR [137-144]. Key to understanding of this proposal, is the fact that the role of these enzymes in activation of p70^{S6k} was determined by effects of two natural products on phosphorylation and enzyme activity: rapamycin, which indirectly inhibits mTOR activity, and wortmannin, which directly inhibits PI(3)k activity. It is also important to note that no direct upstream kinases or other regulatory proteins of p70^{S6k} have been identified to this date.

We will study interactions of p70^{S6k} with its known substrate S6 as a test system for the DHFR PCA in *E. coli* and in mammalian cell lines. We will also seek to identify novel interactions with this enzyme that would lead to new insights into how this important enzyme is regulated. Also, since activation of the enzyme is mediated by multiple pathways that can be selectively inhibited with specific drugs, this is an ideal system to test PCAs as methods to distinguish induced *versus* constitutive protein-protein interactions.

a) Testing of the *E. coli* survival assay: Interaction of p70^{S6k} with S6

This test is ideal, because the apparent K_M (= 250 nM) of p70^{S6k} for S6 protein[145] is approximately the same as the K_D for leucine zipper-forming peptides from GCN4[146] used in our test system. However, we will have to use a constitutively active form of the enzyme for our tests. An *N*-

terminal truncated form of the enzyme D77-p70^{S6k}, is constitutively active and will be used in these studies[147].

Methodology: D77-p70^{S6k}-F[1,2] fusion and D77-p70^{S6k}-F[3] fusion, or F[1,2] and D77-p70^{S6k}-F[3] fusion (as a control) will be cotransformed into *E. coli* and the cells grown in minimal medium in the presence of trimethoprim. Colonies will be selected and expanded for analysis of kinase activity against 40S ribosomal subunits, and for coexpression of the two proteins.

b) Modification of the bacterial survival assay for library screening: Identification of Novel Interacting Proteins. Screening an expression library for interactions with a given target (p70^{S6k}-D77, in this case) will be straightforward in this system, given that the only steps involved are: 1-construction of the fusion-expression library as a fusion with mDHFR fragment[3]; 2-transformation of the library in *E. coli* BL21 harboring pRep4 (for constitutive expression of the *lac* repressor; this is required in the case where a protein product is toxic to the cells) and a plasmid coding for the fusion: p70^{S6k}-D77-[1,2]; 3-plating on minimal medium in the presence of trimethoprim and IPTG; 4-selection of any colonies that grow, propagation and isolation of plasmid DNA, followed by sequencing of DNA inserts; 5-purification of unknown fusion products *via* the hexaHis-tag and sizing on SDS-PAGE.

Methodology: The overall strategy is illustrated in Figure 2. 1-Construction of a directional fusion-expression library: i-cDNA production: We will isolate poly(A)$^+$ RNA from BA/F3 cells (B-lymphoid cells) because these cells have successfully been used in the study of the rapamycin-sensitive p70^{S6k} activation cascade [139]. To enrich for full-length mRNA, we will affinity purify the mRNA via the 5' cap structure by the CAPture

method[148]. Reverse transcription will be primed by a "Linker Primer": it has a poly(T) tail to prime from the poly(A) mRNA tail, and an Xho*I* site for later use in directional subcloning of the fragments. The first strand is then methylated. After second strand synthesis and blunting of the products, "EcoR*I* Adapters" are added, pruducing digestion of the linkers with EcoR*I* and Xho*I* (the inserts are protected by methylation) produces full-length cDNA ready for directional insertion in a vector opened with EcoR*I* and Xho*I* . Because the success of library screening depends largely on the quality of the cDNA produced, we will use the above methods as they have proven to consistently produce high-quality cDNA libraries. ii-Insertion of the cDNA into vectors: The library will be constructed as a *C*-terminal fusion to mDHFR F[3] in vector pQE-32 (Qiagen), as we have obtained high levels of expression of mDHFR fusions from this vector in BL21 cells (Detail in accompanying manuscript). Three such vectors will be created, differing at their 3' end, which is the novel polycloning site that we engineered (details in accompanying manuscript), carrying either 0, 1, or 2 additional nucleotides. This allows read-through from F[3] into the library fragments in all 3 translational reading frames. The cDNA fragments will be directionally inserted at the EcoR*I* and Xho*I* sites in all three vectors at once. 2, 3, 4, and 5- These steps have been described (Details in accompanying manuscript) apart from the final sequencing of clones identified using sequencing primers specific to vector sequences flanking sites of library insertion. The protein purification will also be as described (Accompanying manuscript), by a one-step purification on Ni-NTA (Qiagen). If the product size is more than 15 kDa over the molecular weight of the DHFR component (equal to a cDNA insert of more than

450bp), we will have the inserts sequenced at the Sheldon Biotechnology Center (McGill University).

c) Development of the Eukaryotic Assay We have undertaken the transformation of the system described above, to produce an equivalent assay for use in eukaryotic cells. The basic principle of the assay is the same: the fragments of mDHFR are fused to associating domains, and domain association is detected by reconstitution of DHFR activity in eukaryotic cells (Figure 2.).

Creation of the expression constructs. The DNA fragments coding for the GCN4-zipper-mDHFR fragment fusions were inserted as one piece into pMT3, a eukaryotic transient expression vector [126]. Expression of the fusion proteins in COS cells was apparent on SDS-PAGE after ^{35}[S]Met labeling.

Survival assays in eukaryotic cells. We are currently testing two systems for detection of mDHFR reassembly, in parallel: i- CHO-DUKX B11 cells (Chinese Hamster Ovary cell line deficient in DHFR activity) are cotransfected with GCN4-zipper-mDHFR fragment fusions. The cells are grown in the absence of nucleotides; only cells carrying reconstituted DHFR will undergo normal cell division and colony formation. ii- Methotrexate (MTX)-resistant mutants of mDHFR F[1,2] are being created, with the goal of transfecting cells that have constitutive DHFR activity such as COS and 293 cells. The substitution of Leu22 by Phe produces enzyme with a 20 x increase in resistance and essentially normal catalytic activity[149]. The Leu22Phemutant will be tested in reconstitution of mDHFR fragments to produce enzyme that can sustain COS or 293 cell growth while under the selective pressure of MTX, which will eliminate background due to activity of the native enzyme. The Leu22Phe mutation offers an advantage in selection while presenting no apparent disadvantage with respect to reassembly of active enzyme. If the

reconstituted mDHFR produced in either of the survival assays allows eukaryotic cell growth that is significantly slower than growth with the wild-type enzyme, thymidylate will be added to the growth medium to partially relieve the selective pressure offered by the lack of nucleotides.

d) Testing of the eukaryotic survival assay It will be necessary at the outset to test whether *induced* interactions with $p70^{S6k}$ can be detected. We will use the same test system as that for the *E. coli* test system described above: Induction of association of $p70^{S6k}$ with S6 protein.

Methodology: mDHFR Leu22Phe mutant S6-F[1,2] and $p70^{S6k}$-F[3], or F[1,2] and $p70^{S6k}$-F[3] (as a control) will be cotransfected into COS cells and the cells will be serum starved for 48 hours followed by replating of cells at low density in serum and MTX. Colonies will be selected and expanded for analysis of kinase activity against 40S ribosomal subunits, and for coexpression of the two proteins. Further controls will be performed for inhibition of protein association with wortmannin and rapamycin.

e) Modification of the eukaryotic survival assay for library screening. An important part of the work required in creating a library for use in eukaryotic cells will have been accomplished already, as the EcoR*I*/Xho*I* directional cDNA produced by the Stratagene "cDNA Synthesis Kit" can directly be inserted directionally into the Stratagene Zap Express vector.

Methodology: Steps 1 through 5 are parallel to those for the bacterial library screening (above). 1-Again, the library will be constructed as a C-terminal fusion to mDHFR F[3]. F[3] (with no stop codon) will be inserted in frame in Zap Express, followed by insertion of the novel polylinkers allowing expression of the inserts in all three reading frames (described above), and by the EcoR*I*/Xho*I* directional cDNA. This bacteriophage library will be propagated and treated with the Stratagene helper phage to excise a

eukaryotic expression phagemid vector (pBK-CMV) carrying the fusion inserts. 2-Cotransfection of the library and $p70^{S6k}$-F[1,2] constructs in eukaryotic cells: we will perform the screening in COS or 293 cells, as these are responsive to serum in activating the $p70^{S6k}$ signaling pathway. Selection experiments will be performed as described for the S6 test system above. 3-Propagation, isolation and sequencing of the insert DNA will be undertaken. 4-The cloned fusion proteins will be sized on SDS-PAGE by direct visualization after ^{35}S-Met/Cys labeling, or by Western blotting using a commercial polyclonal antibody to mDHFR.

EXAMPLE 2

Other Examples of Protein Fragment Complementation Assays

Four other assays are herein examplified. The reason to produce these assays is to provide alternative PCA strategies that would be appropriate for specific protein association problems such as studying equilibrium or kinetic aspects of assembly. Also, it is possible that in certain contexts (for example, specific cell types) or for certain applications, a specific PCA will not work but an alternative one will. Below are brief descriptions of each of the PCAs we are presently developing.

1) The Ubiquitin "Split gene assay" Based on the assay originally developed by Johnsson and Varshavsky in yeast [102], we have adapted this assay to a mammalian cell line, COS-7 cells. Ubiquitin can be split into two complimentary polypeptides. When these peptides are combined to form the native structure, the protein becomes susceptible to cytosolic peptidases that recognize a peptide at the C-terminus of ubiquitin. However, the two fragments will not recombine unless oligomerization domains are introduced into the complementary fragments. This behavior has been demonstrated

with the coiled-coil forming GCN4 "leucine zipper". An induced oligomerization can then be detected by measuring whether a reporter peptide has been cleaved from the ubiquitin by cytosolic ubiquitinases. We are now using this approach to probe the interactions of the receptor for the cytokine erythropoietin (Epo) with intracellular partner proteins. These proteins include the kinase JAK2 and PI3-kinase, the phosphates SH-PTP1, and several STAT (signal transducers-activators of transcription) proteins.

2) Glutathione-S-Transferase (GST) GST from the flat worm *Schistosoma japonicum* is a small (28 kD), monomeric, soluble protein that can be expressed in both prokaryotic and eukaryotic cells. A high resolution crystal structure has been solved[150] and serves as a starting point for design of a PCA. A simple and inexpensive colorimetric assay for GST activity has been developed consisting of the reductive conjugation of reduced glutathione with 1-chloro-2,4-dinitrobenzine (CNDB), a brilliant yellow product[151]. We have designed a PCA based on similar structural criteria used to develop the DHFR PCA using GCN4 leucine zippers as oligomerization domains. Cotransformants of zipper-GST-fragment fusions are expressed in *E. coli* on agar plates and colonies are transferred to nitrocellulose paper. Detection of cotransformation is detected in an assay where the nitrocellulose is applied as an aerosol with a glutathione-CDNB reaction mixture and colonies detecting co-expressed fragments of GST are detected as yellow images.

3) Green Fluorescent Protein (GFP) GFP from *Aequorea victoria* is becoming one of the most popular protein markers for gene expression[152]. This is because the small, monomeric 238 amino-acids protein is intrinsically fluorescent due to the presence of an internal chromophore that results from the autocatalytic cyclization of the polypeptide backbone between residues Ser65 and Gly67 and oxidation of the - bond of Tyr66[153]. The GFP

chromophore absorbs light optimally at 395 nm and possess also a second absorption maximum at 470 nm[154, 155]. This bi-specific absorption suggests the existence of two low energy conformers of the chromophore whose relative population depends on local environment of the chromophore[156]. A mutant Ser65Thr that eliminates isomerization (single absorption maximum at 488 nm) results in a 4 to 6 times more intense fluorescence than the wild type is obtain by substituting[157]. Recently the structure of GFP has been solved by two groups[158,159], making it now a candidate for a PCA, which we have begun to develop. As with the GST assay, we are doing all of our initial development in *E. coli* with GCN4 leucine zipper-forming sequences as our oligomerization domains. Direct detection of fluorescence by visual observation under broad spectrum UV light will be used. We will also test this system in COS cells, selecting for co-transfectants using fluorescence activated cell sorting (FACS).

4) Fire Fly Luciferase. Firefly luciferase is a 62 kDa protein which catalyzes oxidation of the heterocyle luciferin. The product posesses one of the highest quantum yields for bioluminescent reactions: one photon is emitted for every oxidized luciferin molecule[160]. The structure of luciferase has recently been solved, allowing for development of a PCA[161]. As with our GST assay, cells are grown on a nitrocellulose matrix[162]. The addition of the luciferin at the surface of the nitrocellulose permits it to diffuse across the cytoplasmic membranes and trigger the photoluminescent reaction. The detection is done immediately on a photographic film. The selection is done by simply identifying the flashing colonies. Luciferase is an ideal model for a PCA: The detection assays are rapid, inexpensive, very sensitive, and utilizes non-radioactive substrate that are available commercially. The substrate of luciferase, luciferin, can diffuse across the cytoplasmic

membrane(under acidic pH), allowing the detection of luciferase in intact cells. This enzyme is currently utilized as a reporter gene in a variety of expression systems. The expression of this protein has been well characterized in bacterial, mammalian, and in plant cells, suggesting that it would provide a versatile PCA.

CONCLUSIONS

Our goals at the outset are to fully develop our PCA strategy, first by expanding on our success with the DHFR PCA to meet the six goals we set at the outset, namely, to develop assays with the capability to:

1) Allow for the detection of protein-protein interactions *in vivo* or *in vitro*.

2) Allow for the detection of protein-protein interactions in appropriate contexts, such as within a specific organism, cell type, cellular compartment, or organelle.

3) Allow for the detection of induced *versus* constitutive protein-protein interactions (such as by a cell growth or inhibitory factor).

4) To be able to distinguish specific-*versus* non-specific protein-protein interactions by controlling the sensitivity of the assay.

5) Allow for the detection of the kinetics of protein assembly in cells.

6) Allow for screening of cDNA libraries for protein-protein interactions.

As a demonstration of the strategy, we will use it to identify novel interactions with the enzyme $p70^{S6k}$, study its' regulation and how separate signaling cascades converge on this enzyme. Success in this project will mean that we have met objectives 1 through 4 and 6 we set out in the introduction to this proposal.

In parallel to this work we will continue the development of the other four assays discussed at the end of proposal. We also have other candidate proteins that may be useful for development at some stage.

The full potential of the PCA strategy will be met when we can also meet objective 5, detection of the kinetics of protein assembly in cells. It is likely that the best assays to address questions of kinetics of protein assembly will be met by the fluorescent protein systems we are developing. I also mentioned in the introduction that we hope the PCA strategy will serve as a general method for functional genomic screening. In the future we will form collaborations with groups at this university and elsewhere, specializing in rapid screening and robotics methods that will be necessary for high throughput screening. Finally, another aspect of the PCA not explicitly discussed here was drug screening. However, in the $p70^{S6k}$ studies I discussed using two natural products, rapamycin and wortmannin in the experimental strategy. Explicitly, one could use PCA strategies to screen for drugs that block specific biochemical pathways in cells allowing for a carefully targeted and controlled method for identifying products that have useful pharmacological properties. We will also in the future want to explore colaborations with academic groups and pharmaceutical industry to develop PCAs for drug screening.

Although the present invention has been described hereinabove by way of preferred embodiments thereof, it can be modified, without departing from the spirit and nature of the subject invention as defined in the appended claims.

REFERENCES

1. Reed, L.J.:Multienzyme Complexes. *Acc. Chem. Res.* **7**, 40-46 (1974).

2. Lander, E.S.:The new genomics - global views of biology. *Science* **274**, 536-539 (1996).

3. Evangelista, C., Lockshon, D. & Fields, S.:The yeast two-hybrid system - prospects for protein linkage maps. *Trends in Cell Biology* **6**, 196-199 (1996).

4. Guarente, L.:Strategies for the identification of interacting proteins. [Review]. *Proc. Natl. Acad. Sci. USA* **90**, 1639-41 (1993).

5. Adams, S.R., Harootunian, A.T., Buechler, Y.J., Taylor, S.S. & Tsien, R.Y.:Fluorescence ratio imaging of cyclic AMP in single cells. *Nature* **349**, 694-7 (1991).

6. Chien, C.T., Bartel, P.L., Sternglanz, R. & Fields, S.:The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl. Acad. Sci. USA* **88**, 9578-82 (1991).

7. Fields, S. & Song, O.:A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-6 (1989).

8. Gyuris, J., Golemis, E., Chertkov, H. & Brent, R.:Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* **75**, 791-803 (1993).

9. Johnsson, N. & Varshavsky, A.:Split ubiquitin as a sensor of protein interactions in vivo. *Proc. Natl. Acad. Sci. USA* **91**, 10340-4 (1994).

10. Volz, K.W., Matthews, D.A., Alden, R.A., Freer, S.T., Hansch, C., Kaufman, B.T. & Kraut, J.:Crystal structure of avian dihydrofolate reductase containing phenyltriazine and NADPH. *J. Biol. Chem.* **257**, 2528-36 (1982).